

Summer
2015

A Standardized Geodatabase form for Cultural Resource Management

CP-6950 GIS CAPSTONE
BRIAN MILNER

GEORGIA INSTITUTE OF TECHNOLOGY

Abstract

The field of archaeology has been absent of a standardized geodatabase system to store, share, analyze and display data collected from archaeological investigation. Data storage and database system vary greatly between cultural resource management firms and research institutions. This lack of standardization has led to the inability of the collected data to be utilized to its full potential. The cultural resource field generates massive amounts of data on archaeological sites each quarter that is rarely seen outside of the company. This inability to collaborate and share data with not only other archaeologists, but researchers from other fields severely limits the understanding we can learn from the data. For my GIS Capstone project, I chose to look at how to design a beginning of a standardized geodatabase model that could easily be implemented by those in the archaeology field. A base work that would hold the basic elements that were needed and most common for the archaeology field. This model should help improve efficiency of data management as well as help facilitate further analysis of the information collected.

Literature Review

While archaeologists have an ethical obligation and legal obligation to record, maintain, and disseminate the information acquired during their investigations, they often fall short of these goals due to the lack of a standardized database structure. This deficiency can be expressed by several causes. One major source is the lack of literature and training

pertaining to this essential component in the field of Archaeology. A second issue plaguing the field is the sheer expanse of varying methods in which the data is recovered and recorded. There is a severe need for a standardized database system in the field that can serve to incorporate the data generated through the archaeological process as well as integrate that same data with data from other fields. Through the use of relational databases, archaeologists are better able to fulfil their ethical and legal requirements in regards to collection, maintenance and dissemination.

Problems With The Current Trends In Archaeology

In the field of archaeology there is a trend that seems to be growing. That is the lack of literature on the subject of databases in archaeology (Tennant 2007; Keller 2009). Despite the history and success of databases, searching for the subject pertaining the field of archaeology is tantamount to the scenery of the Serengeti Desert. A vast and barren wasteland, devoid of a refuge of knowledge. Occasionally an oasis of information of on the subject can be found that tries to rectify the abysmal state of affair (Kintigh 2006; Tennant 2007; Keller 2009; Anderson et al 2011). The trending focus seems to be on newest and greatest forms of predictive models and visualization techniques (Tennant 2007). Keller (2009) seems to pin down the underlying reason for this missing section of literature when she states "Databases are not sexy. Along with the missing literature there also exists a missing curriculum offered by anthropology departments on the subject. Most

archaeologists are unfamiliar with the processes underlying database integration (Keller 2009). These two things combine together to leave the field weaker as a whole.

As archaeologists we have an ethical obligation to curate and maintain a detailed record in a way that can ensure the long-term existence and accessibility of the information. These are set forth in the "Society for American Archaeology Principles of Archaeological Ethics" (Society for American Archaeology 1996). Datasets should be accessible by multiple users, incorporate and integrate different types of data from different disciplines, increase the ease of future updating, and contain accompanying data about the collection techniques, sources, etc. of the information contained within (Tennant 2007). Every year vast amounts of data are generated around the world by research institutions, private contractors, and amateur enthusiasts. The average cultural resource management firm averages 10 gigabytes worth of data per quarter (Anderson et al 2011). The cultural resource industry produces the enormous majority of information in the field. This information is difficult to access and is often termed "Grey Literature" (Anderson et al 2011), because it rarely is accessed by those outside the firm that produces them. Currently different organizations wind up storing their data in a plethora of ad hoc management systems (Keller 2009). These sources usually only held together with loose thread of paper work and spread sheets. (Anderson et al 2011). Spreadsheets are common place in the field. Legacy databases without integration to other data sets only serve a limited use and hinder the discipline as a whole. While they are certainly useful tools for the analysis of

budgets and sales, they are extremely error prone and not ideal for the manipulation of the cultural data produced by archaeologists (Keller 2009). Without a standard structure in place it is difficult to access many of these key sources of data. This lack of data and the ability to compare and contrast over the varying sources has led to a crippling of the research (Kintigh 2006). This lack of standardization is counterproductive and detrimental to the research that could be possible. "To achieve the potential in the field over the long term there is a pressing need for an information infrastructure that will allow for the archival, access, integration and data mining of disparate data sets" (Kintigh 2006). This can be accomplished with the use of well-structured database to allow for the increasing number of sources of data to be compatible and usable by the community (Keller 2009).

Why Relational Databases?

For the archaeological field, relational databases are considered the strong candidate model (Kintigh 2006; Tennant 2007; Keller 2009). Relational databases have been around since the early 1970's (Lake 2013). Developed by Edgar F. Cobb in the late 1960's and early 1970's, they have been incorporated through varying discipline to great success. A relational database incorporates multiple files of related information into tables and columns. A link between separate files is established using a matching field contained in both that allows for the querying of information that would normally be found in separate places. Another alternative suggested is that of an object-oriented relational model (Anderson et al 2011), which is similar to relational databases. The use of a relational

database over traditional spreadsheet methods hosts several benefits. Logical structuring of the data provides security over traditional systems (Keller 2009). Through the use of domains available in the relational geodatabase data entry can be more efficiently handled, saving time, and eliminating errors found in traditional recording systems (Tennant 2007; Keller 2009). In archaeology there is no formal set of standards about what is recorded about the specific attributes of the artifacts, there is an underlying structure that is accepted (Anderson et al 2011). The structuring of the relational database can take advantage of this and help to integrate the data. Redundancy is a common practice in everyday archaeological practice. In traditional paper format recording that is normal in the field while working on an archaeological site the same information will be recorded in multiple places and in multiple forms to guard against the inevitable human mistake that are bound to occur. When transferring this idea to an electronic media however, it normally results in confusion and increased time spent editing data, and longer query times (Tennant 2007; Keller 2009). By using a relational database redundancy can be eliminated, which results in more efficient and flexible sets of data than were available using traditional systems (Keller 2009). Another value of the relational database is the inherent inclusion of metadata. Metadata holds information on how databases are formed, scaled and how the differing fields of data relate to one another. They also contain information pertaining to the semantics of the data. This includes how things are recorded, what units of measurement they are recorded in, what spatial context the associated record corresponds

to. It is through this information that it is possible to compare data from different sources and determine how best to glean the results from them (Kintigh 2006).

Database Creation

An essential question for designing a database for any purpose is where to begin? Keller (2009) suggests that for archaeology at least that we begin by creating a striped down view of the pieces of information we want to record. To do this “we must create a data model” (Keller 2009). To do this an entity relationship diagram is useful (Keller 2009; Anderson et al 2011). This technique was developed by Peter Chen in 1976. Entities are dominate feature that we want to collect information about. For archaeological purposes these can be projects, artifacts, shovel tests and excavation units, context, lot, etc. (Keller 2009; Anderson et al 2011). These entities eventually become their own or multiple tables of data that are joined together through a relationship to each other. These relations mainly consist of a one to many relationship. While the types of attributes recorded vary dramatically from project to project and research institution to private firm, there are common types of information that is shared through the archaeological community (Tennant 2007; Keller 2009; Anderson et al. 2011). Anderson et al. (2011) lays out the foundation of an archaeological database by listing six entities that are common through the archaeological community. These include the project, site/area survey, survey unit, context, lot/bag check list, and finally catalog item. Other relationship types are many to many and very rarely one to one (Keller 2009). Keller (2009) suggests that after the overall

entity relationship diagram is made that the attributes and values of the main entities can be addressed. It can be difficult to decide which kinds of data what data needs to be deconstructed into separate but linked tables, what to remove and what to keep. To help in this aspect, 2 principles are employed. The first principle is “Apples and Oranges” (Keller 2009). Data that is dissimilar to each other should be contained in separate tables. Entity data should be in separate tables to each other and not stored within the same table. The second principle is “Asked & Answered” (Keller 2009). The objective here is to no ask the same question of the data multiple times. Also, if there are multiple answers to a question, they should be in their own table (Keller 2009).

Conclusion

There is a void in the knowledge of database standardization and construction in the literature and curriculum of archaeology (Keller 2009). There is currently a pressing need for a standardized database structure in the archaeological community (Kintigh 2006; Keller 2009). This lack leads to data that is not living up to its full potential. Data is recorded in a plethora of formats and regimes. This makes the integration of data from other fields, as well as other archaeological studies, a daunting task. In the current setting it is difficult to obtain data first hand. This means that research must be based on the studies of others second-hand. This leads to an incomplete picture of the broader context of human cultural influence that could be reached with the accompaniment of other fields of science. Through the use of a standardized relational database structure the process of data entry can be

streamlined. Errors can be prevented through a control structure that is missing in the standard spreadsheet data sets that prevalent throughout the field. Incorporation of legacy datasets adds to the knowledge that can be access by more than a handful of people, which sets the stage for a clearer picture of the spatial settings to be discovered.

Currently in the field of archaeology there is not standardized geodatabase system in place. Data collection and management of the spatial information recovered from archaeological surveys and investigations vary from cultural resource firm to cultural resource firm, and research institution to research institution. This lack of standardization is a detriment to the main purpose of archaeology. It causes an increase in difficulty in synthesizing information collected from multiple projects and areas for understanding human history and environmental interactions. Much of the information collected in the field of archaeology comes from private cultural resource management firms. This data is often lost to others in the profession outside of the specific company. This "grey literature" as it is known, has the ability if properly contained to greatly increase the ability of an archaeologist to study regional patterning quickly and to a greater degree than in the current system. In 2006, Keith Kintigh's, article "The Promise and Challenge of Archaeological Data Integration", signals a clear call that the archaeological community desires an integrated and standardized system to aid in the analysis of data. Along with increased efficiency of collaboration within the field of archaeology, a standardized system offers the opportunity to utilize data from other fields of science as well as allowing those same fields the opportunity to utilize archaeological data to further research designs into new areas of better understanding. This paper is written to show how the beginnings of a standardized geodatabase system can be created and implemented quickly and easily by

either a private firm or research facility, as well as providing enough flexibility to incorporate data specific to certain projects or research design. The National Resource Conservation Service maintains a geodatabase of the properties for the National Register of Historical Places. This geodatabase offers a good example of how a geodatabase of archaeological resources could be constructed and maintained.

To begin the project I first needed to understand what data do archaeologists collect and how should it be represented. Through discussions with several professional archaeologists as well as having experience in the field, I determined that there were nine areas of information that made up a basic set of information that both private and public practitioners collected and needed access to on a regular basis. This set consisted of: archaeological sites, archaeological structures, artifacts, archaeological site datums, shovel test pits, test units, features, survey areas and projects. While this list is not a comprehensive list of all the features that are examined and documented in an archaeological investigation, they give a solid beginning point for a standardized geodatabase. Each will be defined briefly.

Projects

I will start by examining the aspect of projects. In the archaeology field a project is the vehicle through which an investigation is enacted. Projects can be as simple as a research objective or as complicated as a well-defined phase investigation. (Anderson et al. 2011) For

each project there certain elements that are collected that lend themselves to a standardized system. These include things like a project name, name and information about the client the project is conducted for, the date in which the project was begun and ended, and who the principle investigator was. Visually a project can be represented with a polygon class that provides a quick reference for the area and scale in which work was conducted.

Survey Areas

Survey area are another important aspect of for archaeology. They indicate the areas in the project areas that are the focus of investigations. They can consist of the entire project area or smaller areas contained within the project area, such as a parcel or tract of land. Survey areas should be represented by a polygon class.

Archaeological Sites

What is an archaeological site? An archaeological site is any place where physical remains of past human activities exist. This is usually established through testing of the area, either through a surface survey or shovel testing. This gives archaeologist an idea of the size of the area in which human activities once occurred. For the database the archaeological site is a representation of the known boundary of the site, indicated by where artifacts are no longer located. There is a large amount of information that has to be collected on a site. Much of it is collected to complete a state site form. Like most data

collected in archaeology, the information for the form varies from state to state. After reviewing site forms for the states of Georgia, South Carolina, Alabama, Tennessee, Florida, and Kentucky, I found that there are however certain aspects that are shared between them. These shared characteristics were broken into accompanying tables that will be discussed further in following sections.

Archeological Structure

For this geodatabase I used the National Register of Historic Places definition of “structure.” “The term "structure" is used to distinguish from buildings those functional constructions made usually for purposes other than creating human shelter.”(National Register Bulletin). For historic buildings used for shelter, I would suggest employing the system used by the National Register.

Datum

A datum is defined as “a fixed, permanent and readily relocatable position used in mapping...” (Neumann and Sanford, 2001). This is important to archaeologists in that it is used to reference all tests, artifacts, and features and their contexts on an archaeological site. Normally these points are arbitrarily chosen and will be given a Cartesian grid coordinate such as North 500, East 500. Corresponding tests will also receive a coordinate related to the datum. A test located 30 meters north of the datum would be referred to as North 530, East 500. This system does not have to be associated with a magnetic northing

alignment and can be used for a “Grid North”. This is usually done when a survey through a survey area is done at a degree other than North/South. An example of this is that of a survey that is following a pipeline corridor that is running at an angle of 75 degrees. In that instance “grid north” would be considered 75 degrees with 255 degrees as grid south, 165 degrees would be grid east and 345 degrees for grid west.

Shovel Test Pits

A shovel test is a small hole dug to a prescribed depth by an archaeologist to determine if there is a presence of archaeological material. The size and depth of a shovel test pit can vary from state to state according to the particular requirements of that state’s historic preservation officer. Usually they are also spaced a set distance apart from each other. This distance is also set by the state historic preservation office and can vary from state to state. Shovel tests are normally used during a Phase I investigation.

Test Units

An archaeological test unit is a larger excavation than a shovel test. Usually the smallest size for a test unit is one meter by one meter. Multiple units can be located adjacent to each other forming an excavation “block.” These units are used on areas that have been established to contain archaeological material during a Phase II investigation. They offer a larger window into the site for determining which cultural periods, type, site formation, and eligibility for the national register.

Artifacts

Artifacts are objects or materials that were made or utilized by humans. They are used in the analysis to determine: what cultural periods/people are represented at a site, what the purpose of the site was, regional trade patterning, etc. These are usually contained in a surface survey, shovel test pit, test unit or in a feature.

Features

Archaeological features include things like soil stains that indicate where storage pits, garbage dumps, structures, or fences once existed. They can also consist of a grouping of artifacts in local area of a test unit.

After establishing the type of features that were normally required for archaeology, it was then necessary to determine the best way to represent these different data features. For both the project and survey areas, I felt that they were best represented with a polygon feature class. This gives the archaeologist a quick visual reference of the location and the amount of work required. This is necessary in establishing logistics in necessary personnel, man hours to complete, etc. Archaeological Sites and Structures were given the ability to be represented as either a point or polygon. This was done to facilitate older data sets that may contain a site or structure without an established boundary. For those represented as a polygon, they can have a centroid point exported if it is necessary for statistical analysis. While both the site and structure can be represented by either a point or polygon, only one

representation is allowed for each site or structure. Datum were represented by a point. Shovel test pits were represented by a point due to the relatively small size. Test units were represented with a polygon class. Artifacts and features, like sites and structures, were represented as either a point or a polygon. The decision ultimately depends on the size of the feature or artifacts as to whether a polygon is employed. Mostly for normal applications a point would be used over a polygon. As with archaeological sites and structures only one representation for each artifact or polygon would be used. For this application, the artifact represented by the actual feature class would be those where exact locations were known. The feature classes are displayed in the Universal Transverse Mercator projected system. This system was chosen because of the familiarity that the archaeological community already has with the UTM system. I chose to use the North American Datum 1983. This is a change from the traditional North American Datum 1927, in which older site information was related using maps based on NAD 27 system.

Once it was determine which aspects would be focused on and the best way to display them, it was then determined what information would be collected for each of the features. I began with projects. In the feature class table I created a field for project number, project name, start and end date, principle investigator, project director, client number and description. The project number field was set as a short integer and is a unique number for each project. This would be used as the primary key for feature class and as the foreign key in relating the project information with other features. The project name field

was set as a text field with a maximum of 225 characters. The client number field is a foreign key that corresponds to a separate client table that can contain more information pertaining to that specific client. The description field was set as a text field with a 3000 character limit. This will allow as much or as little information concerning the particular project to be readily available to the archaeologist.

Next for survey areas, a field was set up for a survey area number. This is the table's primary key. Fields for survey area name, survey start and end date were created, along with a survey category field. This field indicates the level of work involved with each area such as, just survey, testing, or full excavation. Finally a description field was also included to contain information about the area such as vegetation, land use, topology, etc.

Archaeological Sites has a multitude of information that is collected about it. For the feature class itself the attribute table contains fields for site id, official site number, temporary site number, project id, survey area id, site name, site category, source type, source accuracy, source date, and recorder. When a site is first encountered during archaeological work it is commonly given a temporary number until an official number can be applied to it by the state site file office. Once an official number is given to a site that number is used to relate the site to other tables in the database. Most states use the Smithsonian Number system consisting of a two digit state code followed by a two letter county code and finally the number of site it is recorded in the particular county. An example would be 38AK0031. The 38 represents South Carolina, AK represents Aiken

County, and the 31 would mean that it was the 31st recorded site in the county. This can be used because there will not be duplication of the number. Some states currently use alternative systems for site number, however, this standard can be applied to all 50 states in the United States as well as their territories and military holdings as a unique Smithsonian number already exists for them. Before I mentioned that each state collects information for a state site form. While no two states are exactly identical in the information required, they do share many common features. After reviewing five different states, I condensed the common elements found in each into accompanying tables.

General Information Table

The first was a general information. This table contains field for site name and number, state, county, coordinate system, zone, northing, easting, owner, address, extent of investigation, and information about the recorder. The extent of investigation used a coded domain. 1 indicated general survey, 2 indicates testing, and 3 indicates excavation.

Locational Table

The next table contains locational information about the site. These include the site number, coordinate system, datum, northing, easting, USGS quad in which the site is located and the scale.

National Register Information

The national register information table contains details about the sites status, eligibility, condition and justification for inclusion on the National Register of Historic Places. Domain classes for status and significance were used. Status is indicating the potential for inclusion on the national register, PE is used for sites that are potential eligible for inclusion on the national register, NE is used for sites that are probably not eligible, and AW is used to indicate that additional work is needed at the site to determine status. Site significance refers to the scale at which it is an important resource. Significance is coded as 1 for state, 2 for local, 3 national, 4 state and local, 5 for state and national, and 6 for all.

Physical Description Table

The physical description table contains information pertaining to the area in which the site is located. This includes height above mean sea level, distance to nearest water source, water source name, Strahler rank, current land use and vegetation cover, land form type, slope percentage and direction, and soil type and classification. The soil type uses a coded domain based on the USGS soil types. A similar domain for the soil classification could be employed.

Site Characteristic Table

The site characteristics table contains information about the extent of the site, depth below surface that artifacts occur, and if features or human remains are present. It also contains information about the site condition, type of site, preservation on the site and a general description field.

Cultural Period Table

The cultural period table handles information about the category of site it is such as prehistoric or historic. The cultural period field uses a domain to help with consistency. Choices include: Paleo-Indian, early archaic, middle archaic, late archaic, early woodland, middle woodland, late woodland, Mississippian, 16th, 17th, 18th, 19th, and 20th century or unknown prehistoric/historic.

Field Methods Table

This table handles information dealing with what type of investigation was performed. It indicates how the boundary of the site was determined, either by shovel test or surface scatter, how many tests were conducted at the site. This includes both shovel test pits and test units. Information about the size of the screen used during the testing is also recorded.

Documentation Table

The documentation table deals with documents that refer to a site outside of the current project. Fields to record the type of document, and the location of the document

have been created. A field to record the organization that maintains the document as well as a field for an associate file number has also been created.

Photographs Table

This table deals with any photos taken of the site, artifacts, features, or units.

Maps Table

This table records information about maps made of the site, unit or feature.

Archaeological structures records a structure id for a primary key. Fields for site number are used to relate it to the site features. Structure type records the specific type such as earthen mound, historic railroad, brick lined well, historic dam, etc. There are also fields for the source, meaning how the feature was recorded such as GPS, or digitized. This is accompanied by a field for source accuracy and the source date. Finally fields for the name of the recorder and a description of the structure were created.

The datum table records datum id which is used as the primary key. Site number associates the datum with a particular site. Datum number, datum name, datum type, grid northing and easting, source type and accuracy and date, as well as the recorder are also recorded.

The shovel test pit feature class holds a unique shovel test id. A field for site name relates each shovel test to the corresponding site. A survey area id is the foreign key used to

relate shovel tests to the particular survey area. This is done because there can be shovel tests for survey areas that do not have any associated sites. A datum id field uses the primary key from the datum feature class. Fields to record the source type, accuracy and date are included. Finally a field to denote the recorder of the shovel test. Shovel tests also have an accompanying table. This table records additional information such as which transect the test occurred on and which number of test it was on the transect. Depth of the shovel test is recorded in centimeters. Presence of artifacts used a domain field of a yes or no. Artifact count records the total number of artifacts present in the test. Soil texture and color are recorded. A field called condition is usually used if a shovel test cannot be dug. It uses a domain field. It is coded so that WA indicates standing water in the area of the shovel test, BR indicates bedrock encountered, ER means the area is eroded, SL denotes that there is excessive slope in the location of the shovel test, DIST indicates the soil is disturbed, and NA means not applicable. The percentage of surface visibility is also noted. This is used to indicate if there was just a surface survey preformed without a test actually being dug. There is a field to indicate if the test was GPSed. A northing and easting field was included to handle shovel tests that were dug as a delineation. This happens when an artifact is found in a shovel test. Additional shovel tests are dug around the positive shovel test at set intervals. The northing and easting correspond to the appropriate site datum for the archaeological site, and is substituted for transect and test number.

The test unit feature class has a field for test unit id. This is used as the primary key. A field for site number uses the official site number to link to the sites class. A datum id field relates the datum features to the test unit. Fields for source type, accuracy and date are also included. The accompanying table contains a lot of information. Relational fields include an entry id, project id, site number, survey area id, test unit id, feature id, and datum id. For this table the extra fields of entry id and test unit id allow for multiple records for a single test unit to be recorded. Fields for the grid northing and easting help relate the test unit to other feature on the site. The excavator, recorder and supervisor's names are recorded. Unit size is recorded in meters. Through the use of the entry id and test unit id fields allow for multiple levels and zones to be recorded for a single test unit. Level starting and ending depth are recorded for each. Soil texture is recorded using the USGS texture types. A field has been created to record the Munsell Soil Color number and associated color. A field indicates if a flotation sample was collected from the particular level of the unit. Artifact presences and count are also associated with the particular level and zone.

The features class contains fields for feature id, shovel test id, test unit id, site number to associate it with other classes. It also contains a field for feature number, as well as source type, accuracy and date. The final field is the name of the recorder. In the associated table it is organized much like the test unit table. There is an extra entry id field, which like in the test unit field allows for multiple records to be associated with a single feature. A field for feature size is included. The excavation type refers to if it was excavated

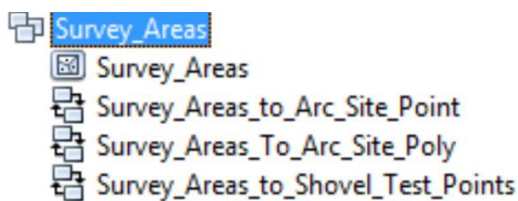
entirely, bisected, or partially removed. The screened field indicates if the feature fill was screened or not. If it was screened the field screen type refers to either dry screening or water screening. Screen size refers to the hole size of the screen used. It uses a domain field that limits the choices to 1/4th inch, 1/8th inch, 1/16th inch, or 1/32th inch. Soil texture, color and Munsell are recorded for inside and outside the feature.

The artifact feature class contains a field for artifact id, site number, shovel test id, test unit id, feature id for the relational aspect. It also contains a field for artifact type as well as a comments field for further information about the artifact. The artifact table contains an entry id, artifact id, site number, shovel test id, test unit id, and feature id field. In addition it contains a field to document the analyst and date which the artifact was examined. A field indicating the date it was excavated. A Pre-Historic/Historic field distinguishes the artifact from pre and post contact periods. The category field is used to define the category of artifact. Example of this are pre-historic ceramic, lithic, historic glass, etc. Artifact type denotes things like debitage for lithics, or cord marked for pre-historic ceramic. Count is for how many of that particular category and type. The relational fields allow for separate entries for multiple artifacts for a specific test unit or shovel test or even specific test unit level and zone.

Now that the feature classes have been created and the tables are set to collect the information how does it all fit together? Using ArcGIS I was able to create relationship classes among the related features and tables. The relationship type that was established

was a one to many relationship. Since there can be many survey areas for one project there is a one-to-many relationship. Survey areas have a one-to-many relationship with shovel test pits and archaeological sites. Shovel tests relate to the survey areas because there can be shovel testing in a survey area that does not produce any sites. Since there are two feature representations for archaeological site (point or polygon) a relationship class has to be established to both. So for example a one-to-many relationship class between survey area to archaeology site point and a separate relationship class between survey area to archaeology site polygon. This must be done anytime there are two possible representations and their related tables and features.

Figure 1: Survey Areas Relationship Classes

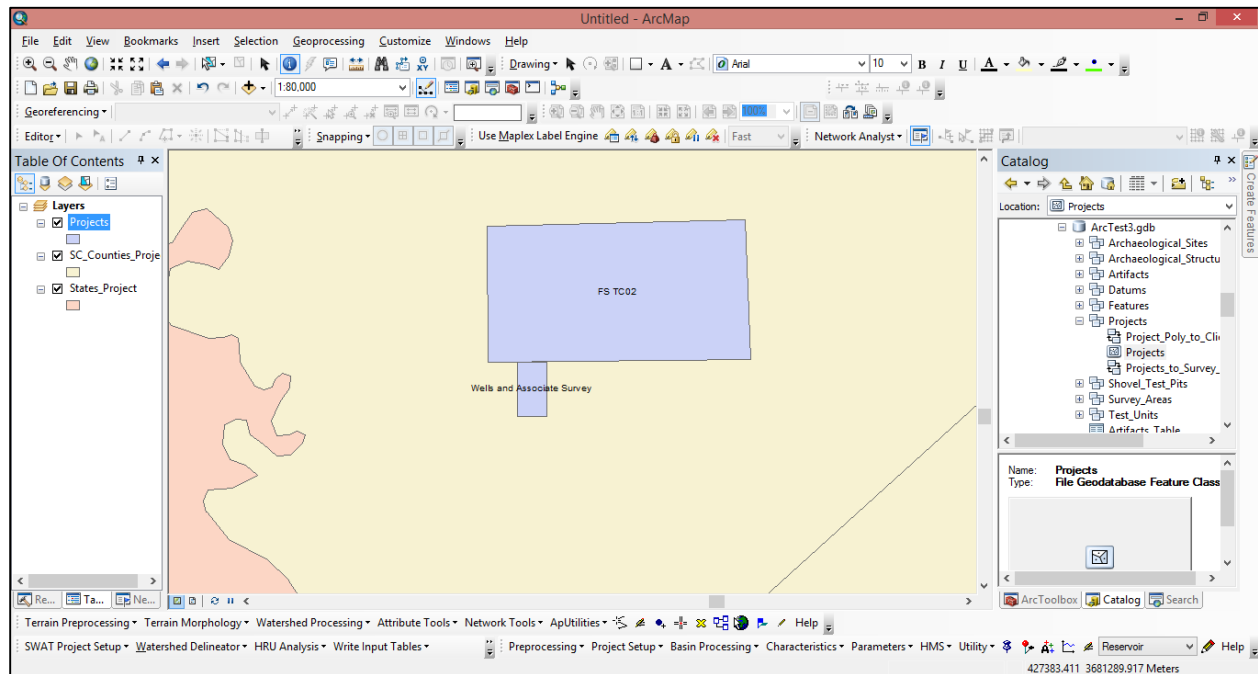


Once the relationship classes were established it was time to enter some data and test the database. I used a combination of provided data and hypothetical data. Site points and polygon locations were provided by the Savannah River Archaeological Research Program from their data sets. Some included extra information such as if the site was pre-historic, historic or multi-component, or lithic scatter, or historic home site. For data to test the capabilities of the other features such as projects, survey areas, shovel test, I created my own to test that the relationship classes were indeed working. I created unique projects.

The survey area for the first project was the same as the project boundary. A total of 555 shovel tests were created in the survey area. These were then associated with an archaeological site (38AK0892) from the provided data, if they occurred within the site boundary. Of those 293 that occurred inside the site boundary were listed as positive for artifacts. Four test units were also created inside the boundary for 38AK0892. Each test was given 10 different level and zone records in the test unit table. Random levels were assigned positive or negative for the presence of artifacts. For the second project, 17 separate survey areas were created inside of the project area. Shovel tests were created in each survey area for a total of 150. Four were made positive and associated with an archaeology site point (38AK0118). An archaeological structure was also created to represent a building associated with the site. For each shovel test that was considered positive, records were created in the artifact table to correspond to the number of artifacts recovered in the shovel test and related back to the shovel test. For test units, artifacts were associated with the specific level and zone they were supposed to have occurred in. This allowed to test the relationship at multiple levels instead of at one level as with the shovel tests.

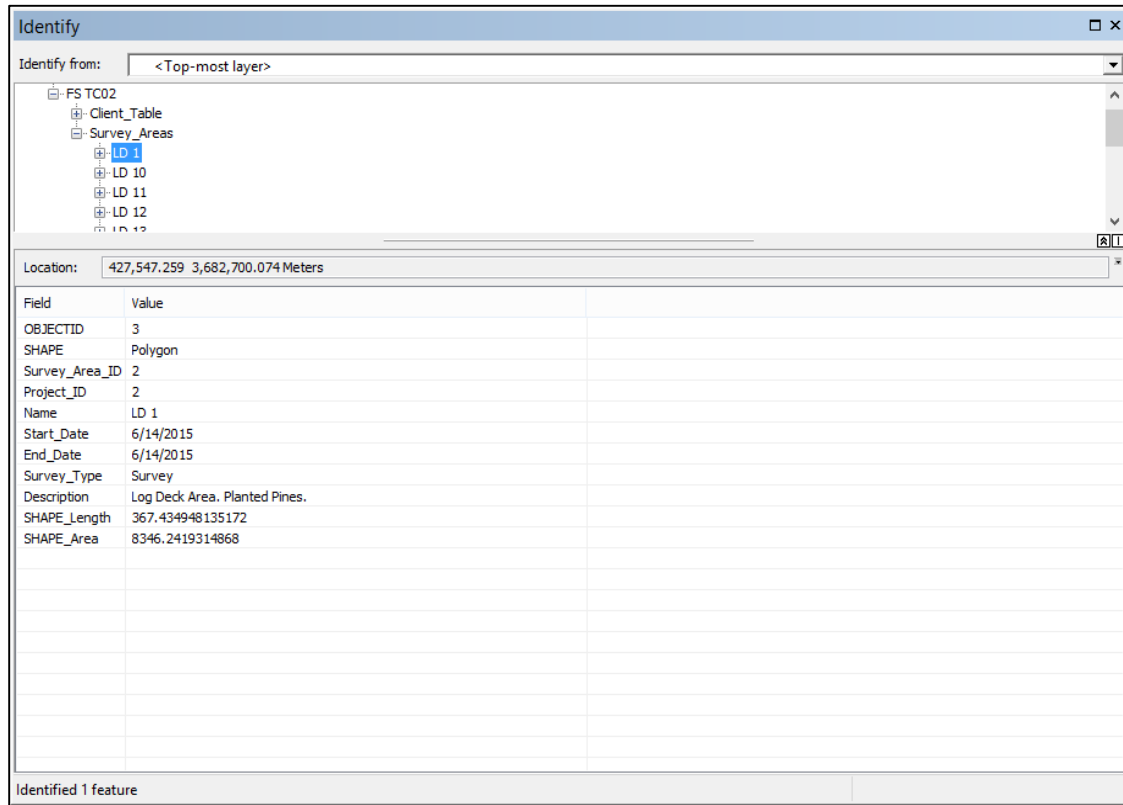
After the relationship classes were set and the data was entered, what was I able to do with the standardized geodatabase that had been created? To begin, I was able to quickly and easily able to identify project areas as shown in figure 2.

Figure 2: Project Locations



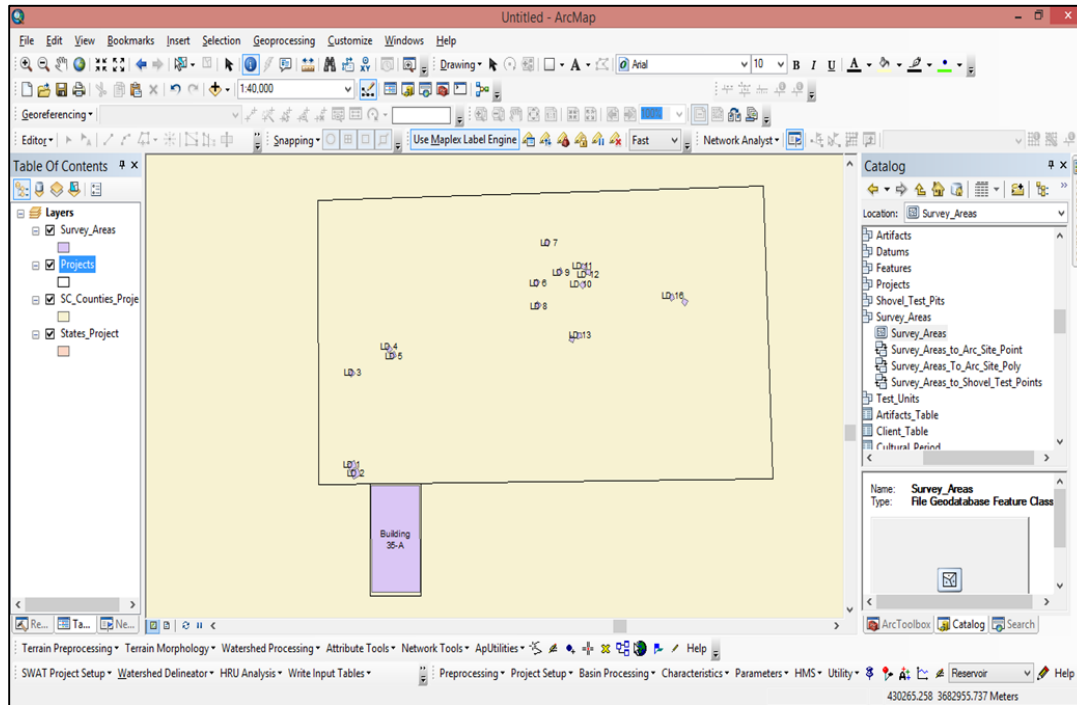
By using the identify tool on the project area you are easily able to access information such as project name that is stored in the project feature attribute table, but are also able to find related information about the survey areas as seen in figure 3.

Figure 3: Identify tool showing related information to the Projects Feature Class from different feature classes.



Along with the project locations, you can see where survey areas are located inside of the project areas. Coupled with the number of survey areas you can quickly total how much area was covered by the surveys. Survey area for Wells and Associates Survey project covered 89% of the total project area, while the seventeen log deck areas located in the FS TC02 project area covered less than 1% of the overall area. Figure 4 shows the various survey areas.

Figure 4: Survey Areas for Two Separate Projects



While in the above example they are displayed as the same color, you could easily go to the symbology setting and have those areas associated with different projects in different colors, see Figure 5. Like with the projects you can use the identify tool to find more information about survey areas, such as number of shovel tests or associated sites as well as the information with them. This can be seen in Figure 6.

Figure 5: Survey Areas Color Coded to Projects

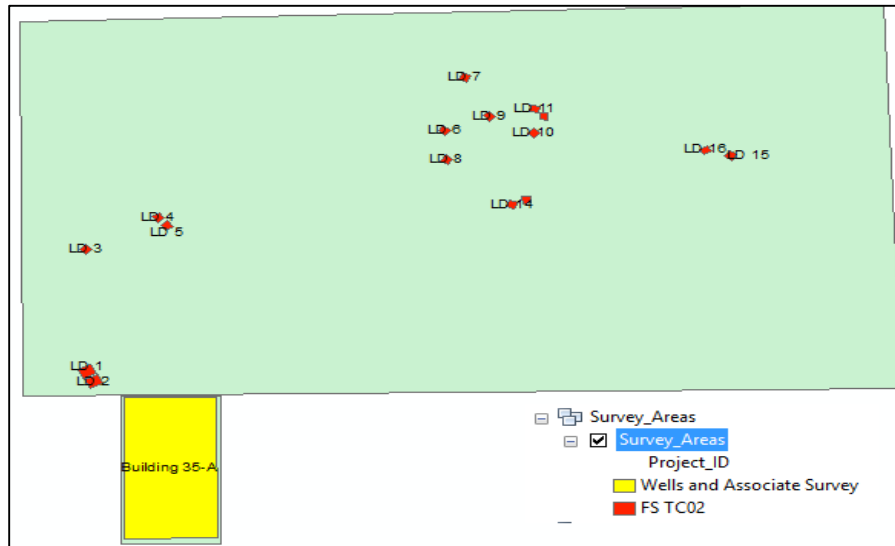
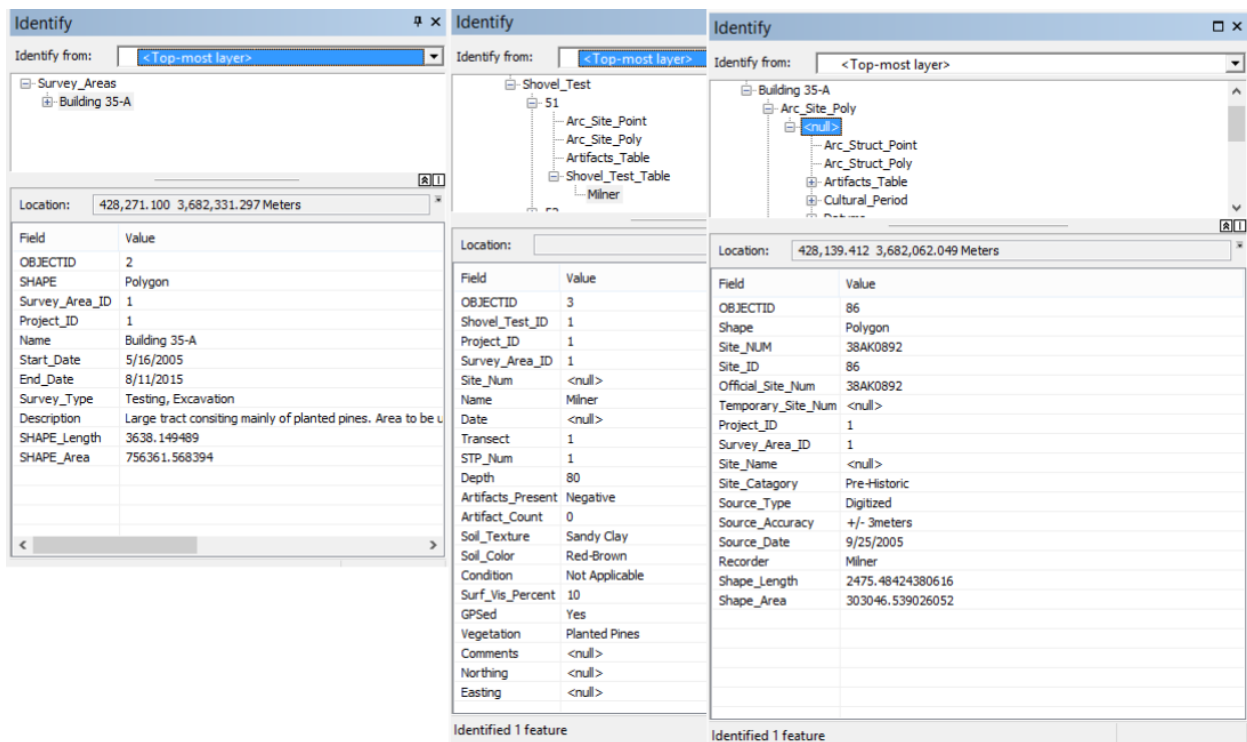
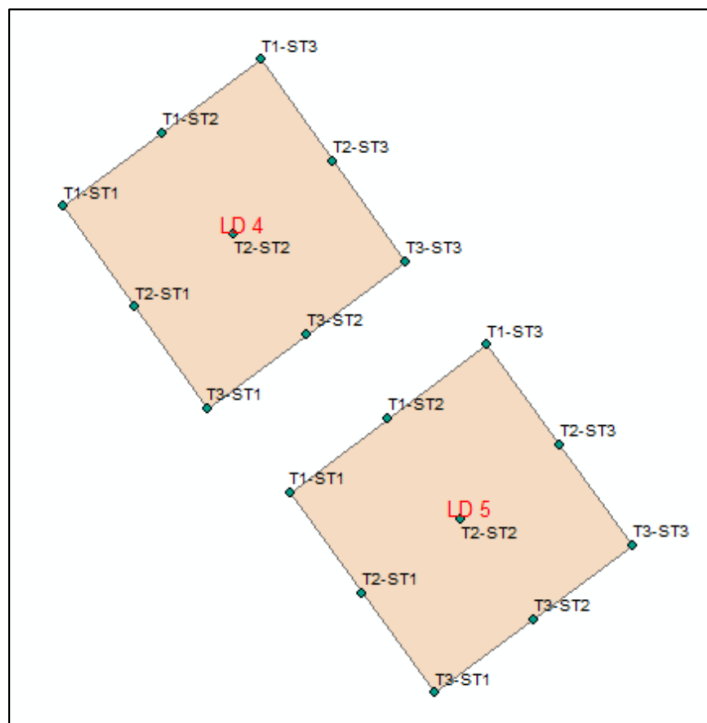


Figure 6: Identify Tool Used on Survey Area.



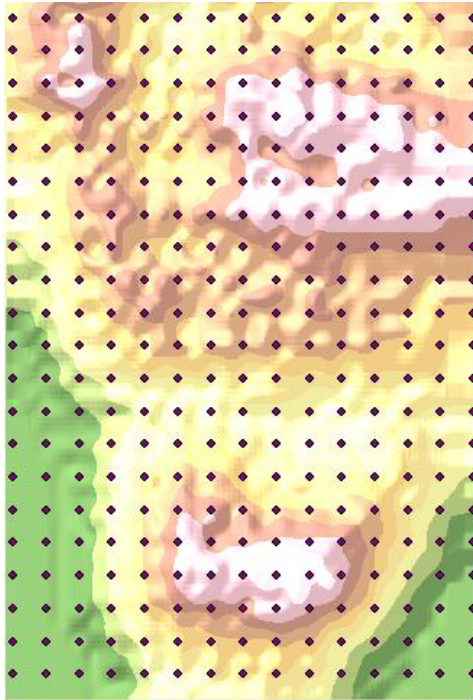
By performing a simple join on the shovel test pits feature class and the shovel test table, you can quickly make a label expression to show which Transect and Shovel Test number each point corresponds to.

Figure 7: Shovel Tests with Transect and Number.



By altering the symbology of the shovel test, the archaeologist can quickly display only the shovel tests that contain artifacts. This can be valuable in discerning site patterning. Having the artifact table related to the shovel test offers the ability to quickly create density maps of artifacts for a site. Figure 8 shows a Kriging map of artifact counts for a portion of 38AK0892 with areas of low density displayed in green and progressing to white for highest areas of artifact concentrations.

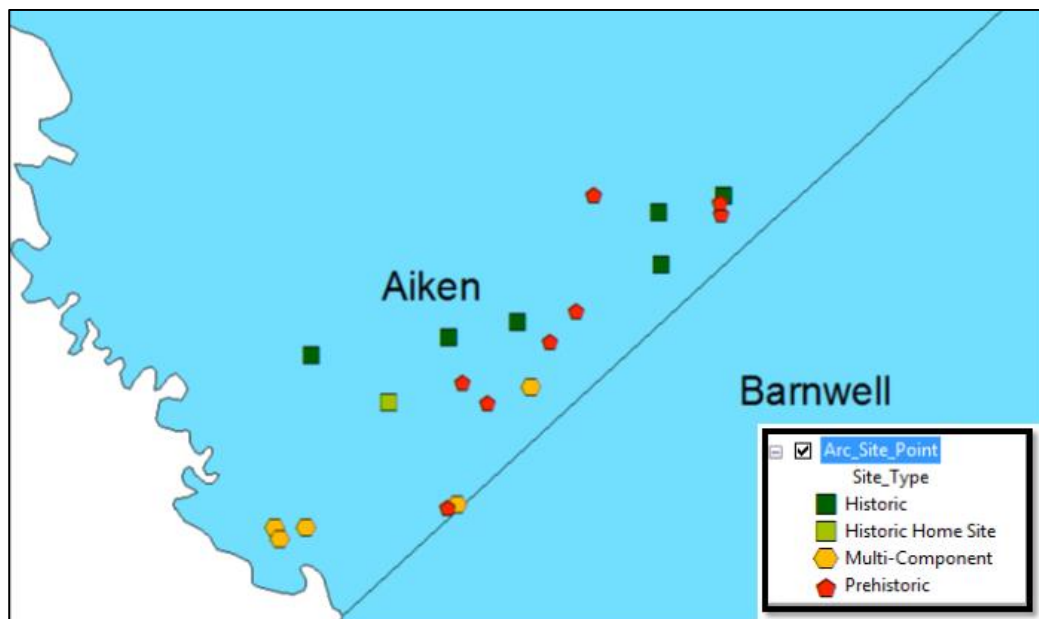
Figure 8: Artifact Density by Artifact Count.



This type of information is invaluable aid in determining the number and placement of test units during Phase II excavation. It also offers insight into site patterning by offering the benefit of being able to quickly switch from artifact count to artifact type. In the current database design, test units are related to their corresponding archaeological site. The test unit table allows for multiple levels and zones to be assigned to a single test unit, much as test unit forms would be filled out in the field. The test units table has also been linked to the artifact table using the particular entry id of each specific level. In this way you can view by not only specific artifacts but by levels over multiple test units for a site. This allows for a chance to analyze occupation zones in depth for sites, offering an even greater detail of intra-site analysis. Through the use of the related tables and the archaeological site feature

class you can quickly and efficiently identify sites as well as complete state site forms. Using the site characteristics table in conjunction with archaeological sites feature class you can quickly differentiate between site types over a local, or regional scale as seen in Figure 9. This type of analysis can be useful in determining settlement patterns. You could easily also examine cultural periods of a site such as Early Woodland vs Mississippian.

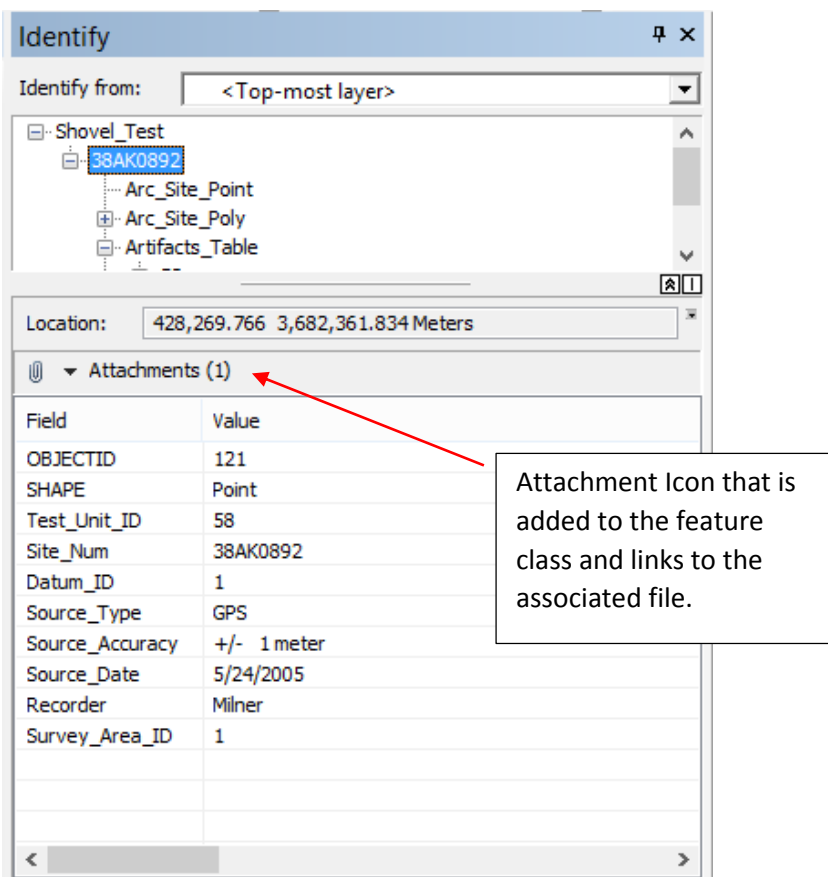
Figure 9: Sites for Aiken County Differentiated by Type.



Another benefit of using an ArcGIS geodatabase is the ability to add attachments to feature classes. This is accomplished by enabling attachment management in ArcCatalog for the particular feature class. When this is enabled, it automatically creates a relationship class and attachment table. Attachments can be added in ArcMap during an edit session through the edit attribute window. Once added an extra attachment button is added to the record

when it is activated using the identify tool. By clicking on the attachment, you are then able to open the file. For this test database, I created an attachment for Shovel Test Pit feature class. Figure 10 shows an example of the attachment in the identify window.

Figure 10: Feature Class Attachments



Overall the design of the geodatabase functioned very well. I was able to access multiple layers of information quickly from the relationship classes that had been established. By simply using the identify tool on a shovel test, I could determine which project it was from, what survey area it was located in, it's transect and number, whether

or not it contained any artifacts, and what site it was located with. The use of primary and foreign keys allows for easy joins between features and tables that allow for more in depth analysis relatively quickly. By using basic classes that are used by both private and public archaeology it facilitates easy use and data sharing between multiple archaeologists. While the current system does not encompass every aspect that could be of interest to researchers, by following the current set up more data could easily be added to the database that is of specific interest to a particular firm or research design. I believe with continuing work and that with further refinements the current system would be more than capable of delivering the needed abilities of data sharing and analysis that is needed in the archaeological field. For online examples of the data and analysis that can be obtained using this system, I have created two ArcGIS Online maps:

1. <http://gtmaps.maps.arcgis.com/home/item.html?id=4fd06d0000b84b9a9ade3fe6ba38e345>
2. <http://gtmaps.maps.arcgis.com/home/item.html?id=61544fe7423141fda1640493cd38d687>

References

- Anderson, David G., Nicholas Hermann, Shane Miller, and Stephen J. Yerka. 2011. "Archaeological Information System Standard Design Concepts." *Digital Data Interest Group Electronic Symposium*. Sacramento: Society for American Archaeology. Accessed June 10, 2015. <http://visiblepast.net/see/americas/archaeological-information-system-standard-design-concepts/>.
- Archaeology, Society for American. 1996. "Society for American Archaeology Principles of Archaeological Ethics." *American Antiquity* 61: 451-452.
- Keller, Angela H. 2009. "In Defense Of The Database." *The SAA Archaeological Record* 9 (5): 26-32.
- Kintigh, Keith. 2006. "The Promise and Challenge of Archaeological Data Integration." *American Antiquity* 71 (3): 567-578. Accessed 06 19, 2015. <http://www.jstor.org/stable/40035365>.
- Lake, Peter and Paul Crowther. 2013. *Concise Guide to Databases: A Practical Introduction*. London: Springer-Verlag.
- Neumann, Thomas W., Robert M. Sanford. 2001. *Cultural Resources Archaeology: An Introduction*. New York: Alta Mira Press.
- Tennant, Edward W. 2007. "A Sample Geodatabase Structure for Managing Archaeological Data and Resources with ArcGIS." *Technical Briefs in Historical Archaeology* 2: 12-23. http://www.sha.org/documents/technical_briefs_articles/vol2article_02.pdf.